EE-0542

April 1983

# A BAYESIAN SPECIFICATION ANALYSIS:
## An Application of Leamer's SEARCH to Air Pollution Aggregate Epidemiology

By


Scott E. Atkinson
University of Wyoming
Laramie, Wyoming 82071

Thomas D. Crocker
University of Wyoming
Laramie, Wyoming 82071

Robert G. Murdock
University of Wyoming
Laramie, Wyoming 82071



USEPA Grant # R808893010

Project Officer:
Dr. Alan Carlin
Officer of Policy Analysis
Office of Policy and Resource Management
U.S. Environmental Protection Agency
Washington, D.C. 20460

OFFICE OF POLICY ANALYSIS
OFFICE OF POLICY AND RESOURCE MANAGEMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

## I. INTRODUCTION

Most decision problems encountered by environmental policymakers involve uncertainty. Given the difficulties of conducting controlled experiments to improve understanding of environmental phenomena, statistical inference is frequently employed. This research usually involves a statistical model that is highly dependent on the investigator's prior beliefs about the relationship between the dependent variable of interest and a list of explanatory variables. When the policymaker uses the results of this research to select a course of action, he employs the combined result of the investigator's data and prior beliefs. Unfortunately, the investigator's prior beliefs are not often reported completely.

This paper focuses on statistical information generated by the Lave and Seskin (1970, 1973, 1977) and the Lave and Chappie (1982) studies of the human health impacts of air pollution. After having admittedly engaged in substantial pretesting, the authors of these studies report a selected set of results. However, they provide little information about the role in selection that their prior beliefs have played; that is, they do not report the robustness of their reported results with respect to key parameters of interest (focus variables) as the set of included explanatory variables (doubtful variables) changes. Since different sets of doubtful variables may be equally plausible *a priori*, the investigator should report the sensitivity of the estimates of the signs and magnitudes of the focus variables to changes in the list of included doubtful variables. A failure to consider and report results for the full range of alternative model specifications which could be "true" means that the opportunity for the policymaker to select whatever mix of possibly "true" specifications best suits his objectives has been censored. The selection of a single model conforming to the investigator's priors can be misleading when several models that differ in their policy implications have some prior creditability. All available information bearing on the robustness and general validity of the alternative models should be provided the policymaker. Because the selectively reported results of Chappie and Lave (1982) and Lave and Seskin (1970, 1973, 1977) have been so widely cited, we apply Leamer's (1978) procedure in this paper to estimate the specification uncertainty of their models.

## II. A BRIEF HISTORY OF AIR POLLUTION AGGREGATE EPIDEMIOLOGY

Although its influence on policy is unclear, the sequence of papers and books produced by Lave and his colleagues on the human health effects of air pollution has been some of the most frequently referenced work in environmental economics over the last two decades. The basic approach has remained that adopted in the path-breaking effort of Lave and Seskin (1970). Using data for 114 U.S. metropolitan areas, they employed single equation ordinary-least-squares methods to regress 1960 total, infant, and disease-specific mortality rates in each of 114 U.S. metropolitan areas upon average ambient sulfate and particulate concentrations, and assorted demographic and socioeconomic variables. They concluded that the total mortality elasticity with respect to ambient sulfates was 0.05; with respect to ambient particulates, this same elasticity was estimated to be 0.04. In a subsequent paper, Lave and Seskin (1973) increased the sample size to 117, introduced some additional air pollution, demographic, and socioeconomic variables, and tested specifications that were nonlinear in the original variables. The conclusions of their 1970 paper were unaltered, however. Finally, the extremely detailed and carefully written Lave and Seskin (1977) book evaluated 1969 as well as 1960 metropolitan area data, employed a variety of cross-sectional, time-series, and pooled models, and yielded nearly the same conclusions.

Recently, Chappie and Lave (1982) have collaborated to reconfirm the results reported in Lave and Seskin (1970, 1973, 1977). They reestimate earlier models with 1974 data for 104 U.S. metropolitan areas. The only new result of importance was the increased sulfate elasticity (now 0.13) and the reduced particulates elasticity (now 0.006). Additional general conformations are provided by several authors who have been inspired to adopt the Lave-Seskin techniques and to apply them to different aggregate epidemiology data sets.[1/]

These confirmations have nevertheless failed to deter numerous critics who, as Chappie and Lave (1982) note, criticize the aggregate nature and the poor quality of the data, and raise issues of omitted variable bias, incorrect functional forms, and the presence of simultaneity. The critics' general procedure has been to use the same or similar data and to find a model which provides air pollution coefficients contradicting the Lave-Seskin results. According to Freeman (1982), Viren (1978) proceeds by adding assorted explanatory variables to the Lave and Seskin (1970, 1973) regressions until a combination is found that reduces the air pollution coefficients to statistical insignificance. Thibodeau, et al. (1980) achieve the same result as Viren (1978) by removing a set of "outliers" from the Lave and Seskin (1970, 1973) data. By positing a reciprocal relationship between mortality incidence and physicians per capita, Gerking and Schulze (1981) obtain statistically significant negative air pollution coefficients, i.e., higher air pollution is associated with declines in mortality incidence. Each of these critics concludes that "... small changes in model specification appear to produce comparatively large changes in implications."[2/] Neither "small" nor "large" is defined, however. Whatever these definitions, the obvious thrust of the critics' stance is that it "... may be unwarranted..."[3/] to employ Lave-Seskin type data and methods to infer a consistent link between air pollution and mortality incidence. No hint is provided the reader about how difficult it would have to be to produce these exceptions before the critics could believe that inferences of a consistent link are warranted.

Lave and Seskin (1977) and Chappie and Lave (1982) have responded in kind to the critics. They busily add and delete numerous combinations of explanatory variables, partition their data sets, and experiment with different functional forms, equation systems, and estimators. The estimates for several alternative specifications are reported. For example, for each choice of a mortality dependent variable and its density function, and for each choice of an equation system and functional form, Chappie and Lave (1982) have a stock of 53 measures which they or their critics consider to be plausible candidates for statistically explaining variations in 1974 metropolitan area mortality incidences. Of the ($2^{53}$) possible inclusion-exclusion combinations of these candidate explanatory variables, 9 ordinary-least-squares single equation regressions are reported in which the unadjusted total mortality rate is the dependent variable. Another 12 similar regressions with the nontraumatic mortality rate as the dependent variable are also reported.[4/] This dependent variable also appears in 2 single equation, generalized-least-squares regressions. Finally, four two-stage-least squares regressions that consider the possible simultaneity between physicians per capita and mortality incidence are reported. Clearly, Chappie and Lave (1982) do not exhaust the number of alternative regressions which might have been reported. Without even having to resort to equation systems, nonlinear forms, or restrictions on coefficient signs and magnitudes, anyone who wishes to obtain a contradictory set of results can most likely find them among the ($2^{53}$) single equation linear model choices.

Neither the practitioners nor the critics of the Lave-Seskin type methods have the means to close the debate; they are unable to provide convincing coverage of the range of plausible models. Both the defenders and the skeptics have been quick to point out that the source of the difficulty lies in the lack of a priori information with which to curb the numerous aspiring models. In Koopman's (1949) terms, the estimation exercise therefore becomes an hypothesis search rather than an hypothesis test. The tests being applied are not independent of the information embodied in the sample. One is looking for hypotheses which best fit the data without being able to specify the alternative hypotheses that might find greater or lesser support. According to whether one is a defender or a skeptic of Lave-Seskin type methods, multiple regression analysis is used to browse for significant or insignificant t-statistics.[5/] With the highly aggregated data the Lave-Seskin methods use, there is little prior knowledge either to guide the search for the model that best fits the data or that best uncovers causal relationships.[6/] In the absence of more information with which to structure models, ending this debate requires a complete and communicable method of model searching and a compact format for reporting the results of the entire search.

In spite of the large number of papers using Lave-Seskin methods, only Smith (1977) and Page and Fellner (1978) supply charts that allow the reader to duplicate their work. The latter employ factor analysis and canonical correlation techniques. Each of these techniques follows a purely mechanical yet communicable statistical format to form scalar indices of groups of variables. One then employs standard hypothesis tests to assess the associations among the groups. However, the mechanical nature of the statistical format makes it difficult to introduce restrictions provided by "true" prior information; moreover, the relationship of the indices to any real phenomenon is frequently unclear.[7/]

Though the Page and Fellner (1978) procedures reduce the temptation to arrive at a "final" form for a model by repeated application of hypothesis tests to the same set of data, they do not obviate it. Smith (1977) chose to apply the Ramsey (1969, 1974) tests for specification error to 32 models he regarded as "fairly representative" of those most often accepted as "final" in the Lave-Seskin type literature of the 1970's. His stated purpose was to ascertain whether the "final" models others had arrived at via the pretesting procedures common to the Lave-Seskin type literature were acceptable on the basis of the Ramsey (1969, 1974) tests for incorrect functional form, omitted independent variables, simultaneity, and heteroscedasticity. His remarks contain a hint of surprise that most of the models performed quite creditably according to the tests. Moreover, Smith's (1977) as well as Page and Fellner's (1978) results are consistent with the Lave and Seskin (1977) estimates of the association of air pollution and human mortality. However, it is unclear how to evaluate the alternative specifications with which Smith (1977) and Page and Fellner (1978) work. The prior beliefs of the researchers who originally specified the alternative "final" models are unknown. One therefore has to accept or to reject each separate model, with its unknown priors embedded.

## III. AN ALTERNATIVE APPROACH TO SPECIFICATION ANALYSIS

With 53 or even as few as 6 or 7 explanatory variables available for use and with a number of alternative functional forms, mortality measures, and density functions for each mortality measure, macroepidemiology researchers have numerous ways to impose their prior beliefs about the impact of air pollution upon human mortality. Though Smith (1977) and Page and Fellner (1978) are mindful of the role that priors have played in reported estimates, their efforts are inherently incapable of assessing the range of priors other investigators might have employed. Leamer's (1978) SEARCH method (Seeking Extreme and Average Regression Coefficient Hypotheses) provides this assessment and portrays it with compact summary statistics. The SEARCH method is fully described elsewhere.[8/] In this section, we try only to convey enough of the flavor of SEARCH to allow the reader to form his own judgments about the informativeness of the inferences that our subsequent air pollution aggregate epidemiology estimates furnish.

In accordance with Leamer and Leonard (1981), consider the following simple linear regression:

$$Y_t = \beta x_t + \gamma_1 z_{1t} + \gamma_2 z_{2t} + \mu_t, \qquad (1)$$

where $Y_t$ is mortality incidence, t indexes a set of T observations, $\mu_t$ is an independently and normally distributed error term with mean zero and unknown variance, $\sigma^2$, and $x_t$ is air pollution. The latter term is a focus variable because it is the center of research concern and will therefore be included in every specification the investigator tests. He wants to know the sign and the magnitude of the unknown parameter, $\beta$. Doubtful variables are the $z_{it}(i=1,2)$, because the prior necessity of their presence in (1) is uncertain. These are the variables whose introduction confronts the researcher with a tradeoff between increasing the bias and reducing the variance of his estimates. In air pollution aggregate epidemiology, physicians per capita, percentage college-educated, and percentage over 65-years old are traditional examples of doubtful variables. Alternatively, if one has a prior belief that percentage over 65-years old obviously belongs in any regression that purportedly explains mortality incidence, he would then be insisting it become a focus variable.

Only 2 doubtful variables are included. It might therefore be feasible to estimate and report the four regression specifications resulting from decisions to include or exclude $z_{1t}$ and/or $z_{2t}$. This would clearly sharpen the reader's judgments about the robustness of the estimates; however, the procedure does not allow the investigator to employ any prior restrictions he suspects might apply to the signs and magnitudes of $\gamma_1$ and $\gamma_2$. Leamer and Leonard (1981) suggest that the investigator employ these priors and thereby enlarge the search. Specifically, they urge him to define a composite variable:

$$w_t(\theta) = z_{1t} + \theta z_{2t} \tag{2}$$

where $\theta$ is a variable which reflects the investigator's priors. For each value of $\theta$ combined with the sample data, there is a unique regression specification, and therefore a different estimate, $\hat{\beta}(\theta)$, for the air pollution coefficient. Because $\theta$ can be continuous over the real line, the set of alternative specifications of (1) need no longer be limited only to the four combinations of $z_{1t}$ and $z_{2t}$ based on their exclusion and/or inclusion. An obvious measure of specification uncertainty is then the difference in the extreme values of $\hat{\beta}(\theta)$. If the interval [$\hat{\beta}$min, $\hat{\beta}$max] is small relative to the sampling uncertainty, or if decisions are insensitive to variations in the values of $\hat{\beta}$ over this interval, then the specification is relatively unambiguous. A large difference between $\hat{\beta}$min and $\hat{\beta}$max implies that specification uncertainty plays a large role relative to sampling uncertainty in the overall uncertainty about the value of the focus coefficient, $\hat{\beta}$. In essence SEARCH evaluates specification uncertainty by searching out the extreme values of $\hat{\beta}$ that occur over all possible covariance matrices.

Leamer (1978) demonstrates that the set of all possible values of $(\gamma_1, \gamma_2)$ generated by varying $\theta$ over the real line is an ellipse of constrained estimates. Each value of $\theta$ represents a different constraint, a different point on the ellipse, and thus a different tradeoff between bias and variance. However, the sample data may make some of these points appear to be extremely unlikely. For example, if $\gamma_1$ is the coefficient for percentage of the population 65 years old or more, a coefficient value which allowed 99.9 percent of the population to exceed this age would be unlikely to appeal to the user of aggregate epidemiology data. The set of points to be considered on the ellipse of constrained estimates can be bounded by defining an $\alpha$ percent ($0 \leq \alpha \leq 100$) sample confidence ellipse.[9/] This point set, which is defined by the intersection of the points in the interior of the locus of constrained estimates and the $\alpha$ percent sample confidence ellipse, represents all possible posterior pairs of $(\hat{\gamma}_1,\hat{\gamma}_2)$ that can result from some prior distribution, given that only sample points lying in the $\alpha$ percent confidence ellipse are to be considered. For each confidence ellipse, minimum and maximum values of $\hat{\beta}(\theta)$ can be generated; that is, one can show how different weights on the prior and the sample distributions cause specification uncertainty to vary. Figure 5.1 in Leamer (1978) is helpful in fixing these ideas.

Leamer (1978) provides a role for the precision of the prior distribution by constructing an "information contract curve" completely analogous to the Edgeworth-Bowley contract curve used in the economic theory of exchange for pairs of consumers. In this case, the sample data, which is analogous to one of the consumers, conveys its information via a likelihood function. The other consumer is a researcher who communicates his information by means of a prior distribution. Leamer's (1978) Figure 5.8 and his surrounding discussion show how this contract curve, which is the locus of tangencies between the conflicting information represented by prior ellipses and sample ellipses, is the locus of informationally efficient points that are jointly preferred by the prior and the data. As with any contract curve, one cannot discriminate among points on it unless more structure is introduced. Thus the distance along the curve can be used as another measure of specification uncertainty. Of course, since the curve is a locus of tangencies between prior and sample ellipses, one could restrict his attention to an interval of the curve lying within some $\alpha$ percent confidence level of the data.

Leamer (1978) shows that more structure with which to choose among points on the contract curve is provided by a measure of the relative precisions of the prior and the sample distributions. For example, if the sample information has low relative variance, one would be more interested in that part of the contract curve closer to the least-squares point. Alternatively, if the prior information is more precise, points on the contract curve in the vicinity of the prior point would be preferred. The difficulty is that the precision of the prior distribution is frequently no more than vaguely known. Leamer (1978) proposes to overcome this difficulty with a procedure which identifies the standard deviation a normally distributed prior must have ("prior sigma") in order to be simultaneously on the contract curve and within a particular confidence ellipse. If, for example, the prior $\sigma$ is very informative and one is dealing, say, with the 95 percent confidence ellipse, he may infer that the contract curve point is quite unlikely, since the prior would have had to be quite small in order to generate it.

The discussion has concentrated upon a single prior; however, Leamer (1978) shows that the same procedures may be extended to linear combinations of focus variables. Thus, when different researchers have quite different combinations of priors, the specification uncertainty inherent in each of the combinations may be fully described.

## IV. AN APPLICATION

After having made the explorations reviewed in Section II, Chappie and Lave (1982, pp. 365,371) conclude that their 1974 data shows that:

> "A strong, consistent, and statistically significant association between sulfates and mortality persists .... When related to the EPA's (1979) estimate of abatement costs, these results support and strengthen the conclusions of Lave and Seskin (1977) that stringent abatement of

sulfur oxides and particulates would produce social benefits (based on health effects alone) greatly exceeding social costs. We regard the evidence for stringent abatement as compelling...."

Ordinary-least-squares regression number 2-5 in Chappie and Lave (1982) embodies nearly all their maintained hypotheses about the relation between mortality and air pollution. Most important, its coefficients for the arithmetic mean air pollution measures are very similar to those in their other reported regressions and thus form the basis for the above-quoted conclusion.

$$
\begin{aligned}
\text{1974 TMR} = {} & \underset{(6.19)}{528.819} - \underset{(-0.57)}{3.043}(\text{MINS}) + \underset{(2.87)}{13.866}(\text{MEANS}) - \underset{(-2.34)}{1.774}(\text{MAXS}) \\
& + \underset{(0.73)}{1.234}(\text{MINP}) - \underset{(-1.19)}{1.008}(\text{MEANP}) + \underset{(1.25)}{0.191}(\text{MAXP}) + \underset{(16.27)}{58.417}(\%65+) \\
& + \underset{(3.21)}{2.412}(\%\text{NW}) - \underset{(-1.39)}{0.009318}(\text{MEDINCM}) + \underset{(1.05)}{18.813}(\text{LOGDENS}) \\
& - \underset{(-1.51)}{26.236}(\text{LOGPOPN}) - \underset{(-4.56)}{10.092}(\%>4\text{YRCOLL})
\end{aligned}
$$

The variables are defined in Table 1. Sample size was 104 metropolitan areas. The numbers below the regression coefficients are t-statistics. With a sample size of 104, the unadjusted $R^2$ for this expression is 0.888. Most of the coefficients are intuitively reasonable in both sign and magnitude, and several achieve high degrees of statistical significance.

We now apply Leamer's (1978) SEARCH procedure to this equation. Initially, we take MEANS to be the only focus variable. All other candidate explanatory variables are doubtful in the sense that we doubt that their coefficients differ from zero or from small numbers. The upper and lower bounds of the estimated coefficient for MEANS are therefore the range of estimates that can be produced by examining all alternative weighted average combinations of the regressions formed by omitting or not omitting each of the

TABLE 1
Definition of Variables*

| | | |
|---|---|---|
| 1974 TMR | -- | The unadjusted 1974 mortality rate per 100,000 population from all causes of death. |
| MINS | -- | Smallest 24-hour sulfate reading in micrograms per cubic meter. |
| MEANS | -- | Arithmetic mean of 24-hour sulfate readings in micrograms per cubic meter. |
| MAXS | -- | Largest 24-hour sulfate reading in micrograms per cubic meter. |
| MINP | -- | Smallest 24-hour total suspended particulate reading in micrograms per cubic meter. |
| MEANP | -- | Arithmetic mean of 24-hour suspended particulate readings in micrograms per cubic meter. |
| MAXP | -- | Largest 24-hour total suspended particulate reading in micrograms per cubic meter. |
| %65+ | -- | Percentage of area population at least 65 years old. |
| %NW | -- | Percentage of nonwhites in area population. |
| MEDINCM | -- | Median income of families in area in dollars. |
| LOGDENS | -- | The logarithm of population density per square mile in the area. |
| LOGPOPN | -- | The logarithm of total population in millions. |
| %>4YRCOLL | -- | Percentage of area population at least 25 years old who are college graduates. |

*All acronyms, definitions, sources, and data are identical to those in Chappie and Lave (1982).

doubtful variables. Thus, the regression results that Chappie and Lave (1982) report, and all results they could have reported, must lie within these bounds.

The upper and lower bounds in Table 2 are the extreme values of the coefficients for MEANS with various levels of the data confidence interval. These correspond to the extreme values within the ellipse of constrained estimates referred to in Section III. At the extreme left of the table are the least-squares estimates. The contract curve traces the value of the coefficient for MEANS along the locus of tangencies between the prior ellipses and the sample ellipses. The t-value of the coefficient for the pooling of the sample and the prior evaluated at a particular point on the contract curve is represented by the posterior-t. The value of the standard deviation of the prior distribution one would have to select to obtain the same point on the contract curve is given by the prior sigma. Specification uncertainty is simply the difference between the upper bound and the lower bound of the MEANS coefficient at the indicated levels of confidence in the data. Sampling uncertainty is defined as 4 times the standard error of the focus variable, which corresponds to a 95 percent confidence interval.

For all values of the data confidence in Table 2, the specification uncertainty exceeds the sampling uncertainty. At the prior (prior sigma = 0), the specification uncertainty exceeds the sampling uncertainty by more than a factor of 5 and the lower bound of the MEANS coefficient is -35.9. Moreover, except for a data confidence of 0.250 or less, the lower bound of the MEANS coefficient is negative throughout. In the absence of guidance as to the relative weights to place on priors versus sample information, these results fail to make a compelling case for a statistically significant association

TABLE 2
Extreme Bounds and Uncertainty Measures for
the Coefficient of Mean Sulfates (MEANS)

Standard error (Sample Sigma) of MEANS = 4.826

| Data confidence | 0.0 | .250 | .500 | .750 | .950 | .990 | 1.000 |
|---|---|---|---|---|---|---|---|
| Upper bound | 13.9 | 27.8 | 30.0 | 32.3 | 36.0 | 38.7 | 70.0 |
| Lower bound | 13.9 | .170 | -1.97 | -4.23 | -7.71 | -10.3 | -35.9 |
| Specification Uncertainty | - | 27.970 | 31.97 | 36.53 | 43.71 | 49.0 | 105.9 |
| Contract curve | 13.9 | 8.11 | 8.13 | 8.23 | 8.48 | 8.73 | 20.2 |
| Posterior t-value | 2.87 | 3.76 | 3.88 | 4.02 | 4.26 | 4.46 | -13.7 |
| Prior Sigma ($\sigma_0$) | $\infty$ | 9.53 | 8.23 | 7.23 | 6.12 | 5.50 | 0.0 |

Sampling Uncertainty = 18.92

between arithmetic mean ambient sulfate concentrations and mortality incidence.

One might justifiably argue that some of the variables we have treated as doubtful while constructing Table 2 should really be focus variables. The addition of these new focus variables could cause the conclusions drawn from Table 2 to be altered. We possess strong priors, for example, that increasing the number of people more that 65-years old, will, cet. par., increase mortality incidence. Most air pollution epidemiologists have strong prior beliefs that total suspended particulates, especially their "fine" particulate versions, have undesirable health impacts. Better education supposedly makes one a more efficient producer of health, while higher income increases the demand for health and also reduces the[8] relative price of health-producing services. The influence these and other priors have upon the upper and lower bounds of the coefficients for MEANS at alternative levels of sample data confidence are presented in Table 3. Although the bounds on the MEANS coefficients are nearly always reduced by these priors, the reduction is very small with the sole exception of the lower bound for %65+. As in Table 2, specification uncertainties continue to exceed the MEANS sampling uncertainty of 18.92 for all levels of data confidence down to 0.250. Similarly, the lower bound of the MEANS coefficient for all priors remains negative down to this same data confidence. The lower bound becomes barely positive if one chooses to give substantial weight to the data rather than to the prior. This exception will hardly be sufficient to convince most people of Chappie and Lave's (1982, p. 365) assertion that it is this data rather than their priors which generate "... a strong, consistent, and statistically significant association ..." between sulfates and mortality. Instead, the range of inferences about

TABLE 3
Extreme Bounds on Mean Sulfates (MEANS)
When MEANS and Another Variable are Focus

| Focus Combination | | Data Confidence | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0.0 | .250 | .500 | .750 | .950 | 1.00 |
| MEANS | U | 13.9 | 27.7 | 29.9 | 32.3 | 35.9 | 68.7 |
| and MEANP | L | 13.9 | .180 | -1.96 | -4.23 | -7.70 | -35.7 |
| MEANS | U | 13.9 | 26.9 | 28.6 | 30.4 | 32.8 | 35.7 |
| and %65+ | L | 13.9 | .403 | -1.54 | -3.51 | -6.25 | -10.6 |
| MEANS | U | 13.9 | 27.8 | 30.0 | 32.3 | 36.0 | 70.0 |
| and %NW | L | 13.9 | .178 | -1.96 | -4.22 | -7.69 | -35.7 |
| MEANS | U | 13.9 | 27.8 | 30.0 | 32.3 | 35.9 | 65.5 |
| and MEDINCM | L | 13.9 | .227 | -1.89 | -4.12 | -7.53 | -31.2 |
| MEANS | U | 13.9 | 27.7 | 29.9 | 32.2 | 35.9 | 69.8 |
| and LOGDENS | L | 13.9 | .360 | -1.74 | -3.95 | -7.34 | -33.2 |
| MEANS | U | 13.9 | 27.8 | 30.0 | 32.3 | 36.0 | 69.2 |
| and LOGPOPN | L | 13.9 | .254 | -1.86 | -4.10 | -7.52 | -33.6 |
| MEANS | U | 13.9 | 27.3 | 29.3 | 31.5 | 34.7 | 51.7 |
| and %>4YRCOLL | L | 13.9 | .187 | -1.96 | -4.23 | -7.70 | -28.5 |

U ≡ extreme upper bound.

L ≡ extreme lower bound.

the impact of air pollution on mortality incidence remains wide under a variety of alternative models.

The high degree of specification uncertainty that the MEANS coefficient exhibits in Tables 2 and 3 could, of course, be due to the aggregate nature of the data being employed. As earlier noted, some of the candidate explanatory variables, such as %65+, are obvious focus variables for any expression intended to explain mortality incidence. If the coefficients for these variables also display so much specification uncertainty that they are uninformative, then one might reasonably conclude that little can be learned from this aggregate epidemiology data set. Table 4 presents the extreme bounds for other focus variables, each in pairwise combination with the focus variable, MEANS. With the sole exception of %65+, the range in the extreme bounds is great. Except for the extreme bounds of %65+ and %>4YRCOLL, the signs of the upper and lower bounds usually differ: however, even for these two variables, specification uncertainty exceeds sampling uncertainty at low levels of data confidence. One might reasonably conclude that there are a large number of explanatory variables not included in this data set that would exhibit no less specification uncertainty than is exhibited by the variables in Table 4.

The preceding discussion is limited to the single equation specifications with mortality incidence as the sole endogenous variable that compose nearly all the published work in air pollution aggregate epidemiology. Chappie and Lave (1982) recognize that simultaneities may exist between mortality and certain of their explanatory variables such as %65+. At the same time they admit that their single equation results could be biased due to the omission of medical care and life-style variables. Perhaps because the plausible

TABLE 4
Extreme Bounds on Other Variables When Mean Sulfates (MEANS) and Other Variables are Focus

| Focus Combination | | Data Confidence | | | | | |
|---|---|---|---|---|---|---|---|
| | | 0.0 | .250 | .500 | .750 | .950 | 1.00 |
| MEANP and MEANS | U | -1.01 | 1.46 | 2.28 | 2.93 | 3.42 | 9.18 |
| | L | -1.01 | -3.41 | -4.17 | -4.77 | -5.22 | -9.34 |
| | | Sampling Uncertainty of MEANP = 6.76 | | | | | |
| %65 and MEANS | U | 58.42 | 64.5 | 66.4 | 67.7 | 68.6 | 70.1 |
| | L | 58.42 | 52.8 | 51.3 | 50.3 | 49.8 | 49.3 |
| | | Sampling Uncertainty of %65+ = 14.40 | | | | | |
| %NW and MEANS | U | 2.41 | 3.98 | 4.46 | 4.84 | 5.12 | 7.29 |
| | L | 2.41 | .732 | .170 | -.286 | -.634 | -5.44 |
| | | Sampling Uncertainty of %NW = 3.01 | | | | | |
| MEDINCM and MEANS | U | -.0093 | .0054 | .0099 | .0134 | .0159 | .0320 |
| | L | -.0093 | -.0254 | -.0308 | -.0351 | -.0385 | -.0795 |
| | | Sampling Uncertainty of MEDINCM = .0268 | | | | | |
| LOGDENS and MEANS | U | 18.81 | 23.7 | 28.5 | 32.2 | 34.9 | 54.6 |
| | L | 18.81 | -8.65 | -14.3 | -18.9 | -22.4 | -70.8 |
| | | Sampling Uncertainty of LOGDENS = 71.67 | | | | | |
| LOGPOPN and MEANS | U | -26.24 | 4.36 | 9.34 | 13.2 | 16.1 | 40.6 |
| | L | -26.24 | -27.7 | -33.2 | -37.6 | -40.9 | -80.2 |
| | | Sampling Uncertainty of LOGPOPN = 69.50 | | | | | |
| %>4YRCOLL and MEANS | U | -10.09 | -7.42 | -6.79 | -6.37 | -6.13 | -5.78 |
| | L | -10.09 | -14.2 | -15.6 | -16.8 | -17.6 | -30.0 |
| | | Sampling Uncertainty of %4YROCOLL = 8.85 | | | | | |

U ≡ extreme upper bound.

L ≡ extreme lower bound.